TABLE OF CONTENTS

This page intentionally left blank

3.2 WATER AND ECOLOGICAL CONDITION INDICATORS

During a 4-hour conference call on November 15, 2005, six peer reviewers (see Appendix 3A) critically reviewed eight draft indicators that EPA proposed to include in the water and/or ecological condition chapters of the ROE Technical Document. These comprised:

- Four new national indicators, proposed for both the water and ecological condition chapters, for which data became available after the July 2005 peer review.
- Two water indicators (one national indicator and one regional indicator) that had were reviewed in July 2005 and substantially revised since then.
- Two new ecological condition indicators that were developed based on recommendations from the July 2005 review.

The charge for this teleconference review is provided in Appendix 3B-2. Preliminary comments, developed individually by reviewers prior to the teleconference, are provided in Appendix 3C-2.

This chapter summarizes discussions, and presents consensus conclusions and recommendations, from the November 2005 teleconference review of these eight indicators. The chapter is divided into nine sections—eight for individual indicators and one covering general comments. The following table shows the reviewers' overall recommendations for these indicators.

Table 3.2-1. Peer Reviewer Recommendations for Water and Ecological Condition Indicators

Indicators	Include with Suggested Modifications	Don't Include Unless Critical Modifications Are Made	Don't Include
Sea level	✔		
Sea surface temperature	✔		
Streambed stability in wadeable streams	✔		
Nitrogen and phosphorus in wadeable streams	✔		
Benthic macroinvertebrates in wadeable streams	✔		
Contaminants in lake fish tissue	✔		
Coastal sediment quality index	✔		
Harmful algal bloom outbreaks in the Gulf of Mexico (Region 4)			✔

3.2.1 SEA LEVEL

Consensus Statements	
Overall recommendation	**Include with modifications.**
Critical modifications	• Relate this indicator to ecological effects (e.g., wetland loss, impacts on coastal communities). Include examples in the write-up of areas affected by sea level changes, perhaps by using case studies. • Discuss the appropriate time scales over which sea level trends can be justifiably interpreted. Emphasize that although the satellite data are more accurate, they are not as useful as the tidal gauge data for distinguishing between true climatic variability and episodic events because the satellite data have been collected for only a short time period. • Consider whether Figure 353.2 adds value to the indicator discussion. If it is included in the ROE 2007, the write-up should explain the data shown (as well as the apparent discrepancies between the data in Figures 353.1 and 353.2) and how it should be interpreted. • Include a more detailed description of the interaction between absolute sea level rise and subsistence to provide additional context for this indicator.
Suggested modifications	• *Sea Level* should be viewed as an ambient condition indicator (affecting coastal communities and ecosystems). • Create a graphic depicting major cities or locations at risk and their respective rates of relative sea level rise (refer to a report that NOAA has compiled).
Other comments	• Rates of present and projected sea level rise could be plotted against measured vertical accretion rates for a number of wetlands to show the range of rates of sea level rise that can be sustained by different coastal wetlands.

The peer reviewers agreed that this indicator should be included in ROE 2007 if the critical changes are made. They felt that the indicator was useful in depicting the long-term effects of global climate change, but emphasized that the effects of sea level rise are location-specific. Much of their discussion focused on the need for EPA to link the indicator with these effects along with the causes for changes in sea level. The reviewers suggested that including a description of the full range of potential effects of sea level change as well as regional examples would provide context for the indicator; areas such as the Florida Everglades, the California Bay-Delta, or the Mississippi Delta could be presented as case studies. These regional examples should rely on past data rather than modeled, forecast data. In addition, the reviewers stressed the need for EPA to distinguish between the time scales of the satellite data and tidal gauge data and to differentiate between the two datasets' ability to detect trends over these time scales. Finally, the reviewers recommended that EPA reexamine Figure 353.2 and determine whether it adds value to the explanation of this indicator.

3.2.2 SEA SURFACE TEMPERATURE

Consensus Statements	
Overall recommendation	**Include with modifications.**
Critical modifications	• Add a long-term dataset to this indicator, and explain that satellite data are more accurate, but they are available for too short a period of time to detect true trends in sea level temperature changes. • Include post-1998 NOAA/NASA data (Heinz Center) data when available. Extend the data from NOAA/NASA (Heinz Center depiction) in Figure 344-1, and place this figure after the figure based on long-term data to ensure it plays a less prominent role. • Include examples of the ecological effects of sea surface temperature changes (e.g., mangrove community changes, northern movement of marine fishes, effects on upwelling and declines of Pacific salmon, coral reef expansion).
Suggested modifications	• Complete a similar analysis for the coolest season of the year to determine whether seasonal minimums have increased. (If this analysis cannot be included in the ROE 2007, EPA should definitely include it in the subsequent ROE.) • *Sea Surface Temperature* should be an ambient condition indicator, rather than an effects indicator.
Other comments	• The satellite data will become more useful if they span a decade or more. Even then, however, this time scale is within the temporal scale of natural variability. • The number of references/citations seems sparse, especially compared to those provided for *Sea Level*.

The peer reviewers agreed that this indicator should be included in ROE 2007 if the critical modifications are made. The largest criticism of this indicator related to the short-term nature of the dataset; in a dataset of this length, temperature trends could be blurred by episodic events such as El Niño. Therefore, the most substantial modification offered by the reviewers was to include a long-term dataset in this indicator. While less accurate than the satellite data, the long-term data would allow for a better trend analysis. The reviewers suggested that EPA exclude Figure 344-1 from the ROE 2007 and include a graphic depicting the long-term temperature data instead. One reviewer stated that, although this indicator should be included in the report, it is not a good ecological indicator for the time period for which the data are available.

The reviewers advised EPA to continue using the satellite data, especially as the record length increases. Furthermore, the Agency should add the post-1998 Heinz Center data to the report. One reviewer suggested that EPA complete a similar analysis on the coolest season of the year to discern whether seasonal minimums, which can have dramatic effects on ecosystems, have increased. The reviewers also felt that EPA could provide additional context for this indicator by including examples of observed ecological effects, such as the effects of sea surface temperature on fish populations. Finally, in nearly all of the pre-meeting comments, the reviewers noted that the size of the graphs presented for this indicator prevented the reader from easily interpreting the scale and comparing the results across regions.

3.2.3 STREAMBED STABILITY IN WADEABLE STREAMS

Consensus Statements	
Overall recommendation	**Include with modifications.**
Critical modifications	• Modify the graphic to convey regional differences and/or stream type (e.g., dammed or free flowing). • Modify the graphic to be more easily understood by identifying proportions of the following: 1) streams with fine sediment sizes and unstable streambeds, 2) streams within the range of bed composition that would be considered stable, and 3) streams with large sediment sizes and overly stable streambeds. • Explain that some streams and regions may naturally exhibit relatively low stability (braided, gravel-bed rivers) or high stability (bedrock-dominated channels). • Improve the write-up by more explicitly describing the methodology used to calculate streambed stability, why the method was chosen, how to appropriately interpret the CDFs, and how these measures may reflect anthropogenic or natural stressors on the system.
Suggested modifications	• Include reference conditions explicitly on the graphic as a horizontal line, or add a descriptive picture to depict what type of streambed a given CDF may represent. • Modify the right vertical axis on the graphic by measuring in thousands of kilometers to shorten the numbers and make the chart more readable.
Other comments	• Present this indicator for review by fluvial geomorphologists (e.g., Wolman, Montgomery, Dietrich, Dunn, Graf). • Alternative presentation to the CDF could be used, such as percent of streams surpassing key thresholds presented by region, much like the IBI graphic presented in the next section. • Regional delineations and conclusions can be made from the data as long as a minimum of 30 to 50 data points are available for that region.

The peer reviewers agreed that this indicator should be included in ROE 2007 if the critical modifications are made. The reviewers felt that while this indicator was important, it would be improved if EPA provided additional information about how the indicator was calculated, presented the data in a regional context, and increased the interpretability of the graphic. The reviewers expressed confusion over the meaning of the CDF; they felt that EPA could clarify it not only by improving the graphic but also by including a descriptive picture to depict what type of streambed a given CDF may represent.

The reviewers discussed the graphic at length and urged EPA to increase its interpretability by including reference conditions and/or by breaking up the data into regions or streambed types. The graphic would be more useful if it conveyed the magnitude or extent of the condition and related it to environmental consequences or anthropogenic/natural stressors. The reviewers also noted that if EPA is not committed to continuing this monitoring program, the indicator will not be useful for evaluating environmental trends.

One of the reviewers suggested that all three freshwater indicators be discussed together, since they are derived from the same dataset. These indicators are prone to the same interpretability obstacles, which could be overcome if EPA provided a more detailed explanation of the methodology used to calculate the indicator and defended the single-sample methodology. This same reviewer also suggested in his pre-meeting comments that a map of sample points be shown (refer to http://www.epa.gov/OWOW/monitoring/wsa/WSAProbabilityDesign.pdf for an example).

3.2.4 NITROGEN AND PHOSPHORUS IN WADEABLE STREAMS

Consensus Statements	
Overall recommendation	**Include with modifications.**
Critical modifications	• Modify the graphics to enhance the interpretability of the CDF by presenting data in a regional context and including reference conditions (or a "rule of thumb"). For example, by region, present the CDF showing a threshold value considered "impacted"; 100 μg/l for total phosphorus was discussed, but both nitrogen and phosphorus thresholds should be thoroughly researched. • Present the N/P ratios graphically, in addition to the CDF. • In the write-up, address the concern about single-sampling of each stream; explain why the number of samples taken is appropriate, and reference other studies that address the use of single samples for regional characterization of macroinvertebrates or peer-reviewed publications based on this dataset.
Suggested modifications	• Consider applying Redfield ratios or a stoichiometric index to enhance the indicator's interpretability.
Other comments	• Regional delineations and conclusions can be made from the data as long as a minimum of 30 to 50 data points are available for that region.

The reviewers agreed that this indicator should be included in ROE 2007 if the critical modifications are made. As with *Streambed Stability in Wadeable* Streams, the reviewers emphasized the need to improve the graphics for this indicator, to explain in more detail how the CDF was calculated, and to present the indicator in a regional context. In addition, they directed EPA to defend the single-sample data collection methodology used and to clearly state that the study design does not allow for inferences to be made for individual streams. The reviewers also were interested in the trends that this indicator may reveal if data collection efforts continue in the future. If EPA is not committed to continuing this monitoring program, the indicator will not be useful for evaluating environmental trends.

3.2.5 BENTHIC MACROINVERTEBRATES IN WADEABLE STREAMS

Consensus Statements	
Overall recommendation	**Include with modifications.**
Critical modifications	• Include both measures of benthic macroinvertebrate health (i.e., MMI and O/E) in the ROE 2007. The write-up should detail the differences between the two measures and how each should be interpreted and applied. • Modify the graphic to enhance understanding, and break the data out into regions (see Figure 1, below). • Expand the reference list and applicable sections of the write-up to include more details on the origin of the data and the methodology used to calculate the indices. *(This general concern applies to all indicators.)*
Suggested modifications	• In the write-up, link the IBI CDFs to causal factors and ecosystem effects, even if only in a descriptive, exemplary way. Oregon, North Carolina, Ohio, and several other states have used O/E to determine causal relationships; these applications should be described.
Other comments	• Include a map of the sample points, as shown at: http://www.epa.gov/OWOW/monitoring/wsa/WSAProbabilityDesign.pdf. • Regional delineations and conclusions can be made from the data as long as a minimum of 30 to 50 data points are available for that region.

The peer reviewers agreed that this indicator should be included in ROE 2007 if the critical modifications are made. All of the reviewers stated that they found it difficult or impossible to determine how the indicator was calculated using the references that EPA provided. They also felt that, like the other two freshwater indicators, this indicator should be presented in a regional context along with reference conditions and should be better linked to ecological effects and conditions.

Upon learning that both the Multi Metric Index (MMI) and observed over expected (O/E) approaches for calculating IBIs were available, the peer reviewers advised EPA to include and fully explain both measures in their presentation of this indicator. They emphasized the need to enhance the interpretability of the graphic, and one reviewer included an example presentation of IBI data on a regional basis in the pre-meeting comments (refer to Figure 1). Finally, the reviewers noted that the indicator in its current form is not capable of demonstrating trends over time; thus, EPA should continue this monitoring program if the indicator is to be useful for evaluating environmental trends.

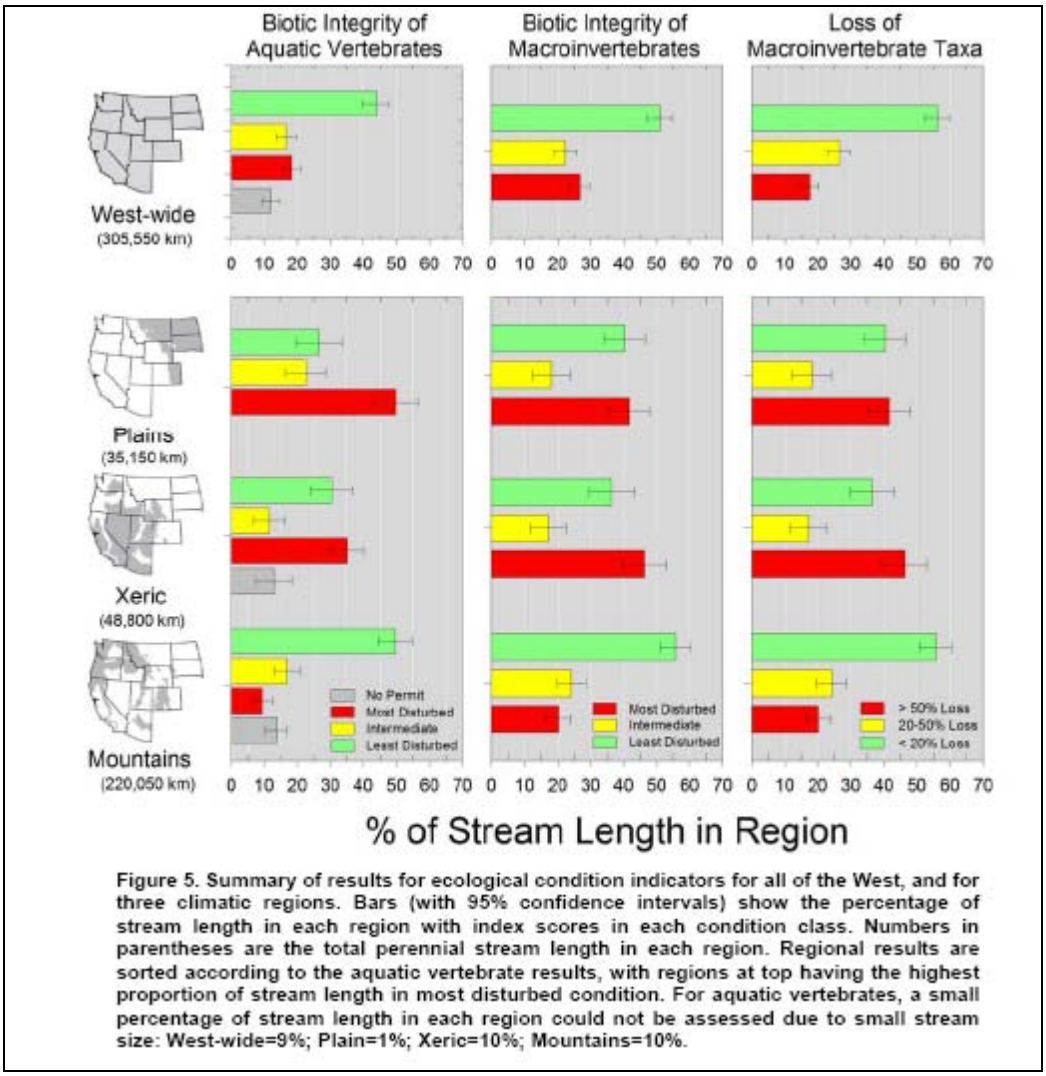

Figure 5. Summary of results for ecological condition indicators for all of the West, and for three climatic regions. Bars (with 95% confidence intervals) show the percentage of stream length in each region with index scores in each condition class. Numbers in parentheses are the total perennial stream length in each region. Regional results are sorted according to the aquatic vertebrate results, with regions at top having the highest proportion of stream length in most disturbed condition. For aquatic vertebrates, a small percentage of stream length in each region could not be assessed due to small stream size: West-wide=9%; Plain=1%; Xeric=10%; Mountains=10%.

Figure 1. Example of Regional Presentation of IBI Data

3.2.6 CONTAMINANTS IN LAKE FISH TISSUE

Consensus Statements	
Overall recommendation	**Include with modifications.**
Critical modifications	• Modify the graphics to include additional descriptions of the information presented, especially in the case of the two apparently identical maps depicting sampling sites. • Include contaminant-specific thresholds and/or criteria in Figures 335-1 and 335-2 to provide context for the results. • In the write-up, emphasize how this indicator should be interpreted with respect to ecological conditions and human health by providing thresholds and/or criteria. • In the write-up, defend the rationale behind collecting different types of fish at different sampling sites (if this indeed was the case).
Suggested modifications	• In the write-up, explain that arid areas are not well represented in the dataset. Because the distribution of sampling sites was based on the frequency of occurrence of lakes, the study cannot address fish tissue contaminants in more arid regions. A different sampling design would be required. • Present the data regionally for those areas where sufficient data have been collected (refer to Figure 2, below). Presenting the data by ecoregions would likely be preferable to EPA regions.
Other comments	None.

The peer reviewers agreed that this indicator should be included in ROE 2007 if the critical modifications are made. Several of the reviewers were unclear about the interpretation of this indicator relative to human and ecological health. Therefore, they suggested improving the indicator graphics by including contaminant-specific threshold and/or criteria levels, such as the mercury fish tissue criterion developed by EPA. In addition, the reviewers noted that there did not seem to be a difference between the two maps of the 500 sampling sites, and they recommended that EPA either delete the second, duplicative map or include two legitimately different maps with appropriate labels and descriptions. One reviewer proposed that the data be divided by ecoregion and presented using a graphic similar to Figure 2, and the same reviewers suggested that the tissue concentration tables could be presented graphically to enable interpretation by a less technical audience.

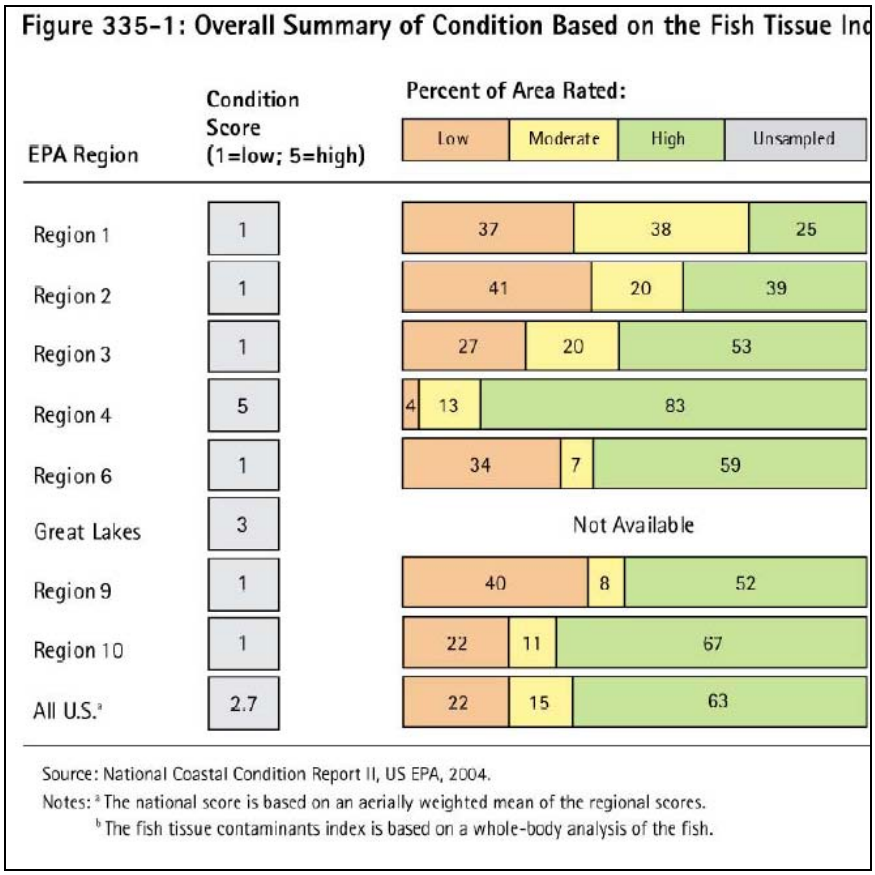

Figure 2. Example of Regional Presentation of Fish Tissue Data

3.2.7 COASTAL SEDIMENT QUALITY INDEX

Consensus Statements	
Overall recommendation	**Include with modifications.**
Critical modifications	• Rename the indicator from "Sediment Quality" to a more appropriate descriptor, such as "Sediment Toxicity," "Sediment Contamination," or a combination of these two terms. • Include additional details about the bioassay methodology, why this particular bioassay and test organism were chosen, and the quality control methods employed when carrying out the bioassay. • Provide a clear explanation of the methodology used to calculate the index. • Clearly and prominently describe the limitation of sediment toxicity tests and the application of the "effects range low" (ERL) and "effects range medium" (ERM) endpoints. • Improve the graphics and write-up to provide an indication of ecological impacts and present trends. Include a more detailed description of the indicator's limitations, and specify appropriate interpretations of the graphics. • Improve the approach for addressing missing data. Explain why the data are missing, and describe any shared characteristics of sites missing data. Exercise care when interpreting the data, as the sampled population to which inferences can be made may differ from the target population.
Suggested modifications	None.
Other comments	• Calculating an index based on bioassays using one species may not be scientifically justifiable or meaningful. • Toxins can produce a wide range of effects including physiological effects, stress, impacts on reproductive processes, disease, and effects on viability (mortality). This range of effects needs to be addressed in addition to mortality when evaluating whether or not sediments exhibit toxicity. • Sediment toxicity tests are most useful and typically applied to a localized area. ERL and ERM endpoints are useful as screening tools but may not be as appropriate as a national or regional indicator. • Even small proportions of estuaries with contaminated sediments can cause widespread effects on biota through the estuary. The text needs to explain the ecological implications of different levels of sediment contamination. Also, one of the regions with the highest sediment contamination exhibited the lowest level of sediment toxicity. The text should help readers understand how this can occur and how the results should be interpreted. • Regional results should be presented accompanied with a map, as few people know where EPA regions are located.

The peer reviewers agreed, with significant reservation, that this indicator could be included in ROE 2007 if all the critical modifications were made. While the reviewers appreciated EPA's revisions to the July version, they did not endorse this indicator to the same degree as those previously discussed. They expressed concern over deriving an index from a single-species bioassay and applying ERL and ERM endpoints without explicitly describing the limitations of sediment toxicity tests and the ERL and ERM. In addition, the reviewers requested that EPA include an expanded description of the procedures used to calculate the index and suggested that the write-up and graphics better link the indicator to ecological effects. One of the reviewers was concerned that the graphic, in its present form, could be misinterpreted if general, regional conclusions were viewed as applying to specific sites within a region. Lastly, the reviewers proposed that the graphics be modified to communicate trends over time.

3.2.8 HARMFUL ALGAL BLOOM OUTBREAKS IN THE GULF OF MEXICO (REGION 4)

Consensus Statements	
Overall recommendation	**Do not include.**
Critical modifications	• Do not include this indicator in ROE 2007, but leave a placeholder for this type of indicator.
Suggested modifications	None.
Other comments	• In future ROEs, include a suite of HAB indicators for several algal groups (e.g., estuarine and coastal dinoflagellates, cyanobacteria, prasinophytes, brown algae), ecosystems, and regions. • The study design could be improved if a researcher delineated a particular area in the Gulf, resampled that area at regular intervals, and counted the number of days per year in which harvesting shellfish from that area was prohibited. This sampling method would allow for a trend analysis over time and would resolve the issue of whether observed red tide trends were the result of the level of sampling rather than a reflection of changes in the frequency or severity of HABs.

The peer reviewers agreed that this indicator should *not* be included in ROE 2007. The reviewers agreed that while HABs are important indicators of ecological health, EPA did not choose the most appropriate regional dataset to represent the indicator. The initiation, location, and transport of red tides in the Gulf of Mexico are generally more attributable to physical forcing mechanisms than anthropogenic influences and are, therefore, less manageable. How would EPA use this indicator in a management context, except possibly as a human exposure warning indicator (which Florida already has)? In addition, the reviewers noted that Figure 237R does not explain the significance of number of blooms versus bloom duration and does not tie either measurement to environmental pressures or ecological effects.

The reviewers agreed that in future ROEs, EPA should include regional HAB datasets for several algal groups and ecosystem types. For example, blue-green algae (cyanobacterial) blooms in freshwater and some brackish systems, as well as dinoflagellates in the Chesapeake Bay and other estuaries, can be linked to nutrient loading and have been mitigated by reducing nutrient inputs. EPA should emphasize these types of HABs, as they represent more manageable environmental stressors and can be sampled under more controlled conditions.

3.2.9 GENERAL COMMENTS

Several universal comments were made throughout the peer review that apply to all or a distinct group of the indicators. Comments that apply to all indicators include:

- EPA should generally consider how it categorizes each indicator into the four broad types of indicators, develop and discuss indicators that relate across all four categories, and explain their categorization for each indicator in its write-up.
- Documentation of the underlying the calculation of all indicators should be easily accessible, and citations should be more precise.
- Graphics should be more readable and made as simple to interpret as possible using appropriate decimal places and clearly visible colors.
- It is important that EPA convey its commitment to continuing the monitoring programs underlying the indicators if they are to be deemed useful for evaluating environmental trends. Degree of certainty of continued measurement should be described for each indicator.

With respect to the three freshwater indicators that were based on the same dataset (i.e., *Streambed Stability in Wadeable Streams*, *Nitrogen and Phosphorus in Wadeable Streams*, and *Benthic Macroinvertebrates in Wadeable Streams*), the reviewers stressed the following:

- The single-sampling method does not allow for inferences to be drawn about a particular stream but does allow for regional conclusions to be drawn as long as a minimum of 30 to 50 data points are available for that region.
- The precision of regional inferences is far more influenced by the number of sites visited than by the precision with which a specific site is evaluated. (Over a wide range of studies, a partitioning of indicator variation demonstrates that site-to-site variation ranges from 98% to 2% of the total variation, with the remainder being mainly residual. For many indicators, the site variation accounts for at least 50% of the total variation. Such evaluations have been factored into the study design through the choice of indicators and the protocols for their evaluation; indicators with primarily residual or local variation have not been included as regional indicators.)
- EPA should explain the common features of these surveys, specifically:
    - The methodology for and rational behind site selection.
    - How to interpret cumulative distribution functions (CDFs) if they remain in the presentation.